

BMJ Open Effects of external inspections on sepsis detection and treatment: a stepped-wedge study with cluster-level randomisation

Gunnar Husabø ^{1,2}, Roy Miodini Nilsen,³ Erik Solligård,^{4,5} Hans Kristian Flaatten,⁶ Kieran Walshe,⁷ Jan C Frich,⁸ Gunnar Tschudi Bondevik,^{2,9} Geir Sverre Braut,^{10,11} Jon Helgeland ¹², Stig Harthug,^{13,14} Einar Hovlid ^{1,2,11}

To cite: Husabø G, Nilsen RM, Solligård E, *et al.* Effects of external inspections on sepsis detection and treatment: a stepped-wedge study with cluster-level randomisation. *BMJ Open* 2020;**10**:e037715. doi:10.1136/bmjopen-2020-037715

► Prepublication history and additional material for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-037715>).

Received 14 February 2020

Revised 17 August 2020

Accepted 10 September 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Einar Hovlid;
einar.hovlid@hvl.no

ABSTRACT

Objective To evaluate the effects of external inspections on (1) hospital emergency departments' clinical processes for detecting and treating sepsis and (2) length of hospital stay and 30-day mortality.

Design Incomplete cluster-randomised stepped-wedge design using data from patient records and patient registries. We compared care processes and patient outcomes before and after the intervention using regression analysis.

Setting Nationwide inspections of sepsis care in emergency departments in Norwegian hospitals.

Participants 7407 patients presenting to hospital emergency departments with sepsis.

Intervention External inspections of sepsis detection and treatment led by a public supervisory institution.

Main outcome measures Process measures for sepsis diagnostics and treatment, length of hospital stay and 30-day all-cause mortality.

Results After the inspections, there were significant improvements in the proportions of patients examined by a physician within the time frame set in triage (OR 1.28, 95% CI 1.07 to 1.53), undergoing a complete set of vital measurements within 1 hour (OR 1.78, 95% CI 1.10 to 2.87), having lactate measured within 1 hour (OR 2.75, 95% CI 1.83 to 4.15), having an adequate observation regimen (OR 2.20, 95% CI 1.51 to 3.20) and receiving antibiotics within 1 hour (OR 2.16, 95% CI 1.83 to 2.55). There was also significant reduction in mortality and length of stay, but these findings were no longer significant when controlling for time.

Conclusions External inspections were associated with improvement of sepsis detection and treatment. These findings suggest that policy-makers and regulatory agencies should prioritise assessing the effects of their inspections and pay attention to the mechanisms by which the inspections might contribute to improve care for patients.

Trial registration NCT02747121.

INTRODUCTION

External assessment of healthcare providers is in widespread use as a policy strategy to foster improvement in the quality of care.¹

Strengths and limitations of this study

- This is the first large-scale study using a robust design to evaluate the effects of external inspections on clinical care.
- As it was not possible to design a randomised controlled study, we used a stepped-wedge design, allowing the inspections to proceed as usual while we assessed effects based on data collected by the inspectors.
- Even though we adjusted for a range of known confounders, there is a risk that unknown external factors not included in the analyses introduced bias to the effect estimates.

WHO defines assessment as an external institutional strategy and divides it into three subcategories: accreditation, certification and supervision.² According to WHO, *accreditation* generally refers to external assessment of an organisation by an accreditation body, *certification* is usually used to describe external assessment of compliance with standards published by the International Organisation for Standardisation (ISO) and *supervision* refers to an authoritative monitoring of healthcare providers' compliance with minimum standards often set by legislation.²

These assessment schemes represent heterogeneous, complex processes that consist of a set of activities that are introduced into varying organisational and regulatory contexts, and their origin and objectives can differ.³ They share an important defining element in that: "some dimensions or characteristics of a health care provider organisation and its activities are assessed or analysed against a framework of ideas, knowledge, or measures derived or developed outside that organisation".⁴ The phrase "external" also implies that the assessment is initiated and

conducted by an organisation external to the one being assessed.

External assessments can serve different purposes. They can represent a control strategy, emphasising whether providers meet certain standards, thereby promoting accountability and transparency in a regulated society.⁵ However, they can also represent an improvement strategy, based on the assumption that externally promoted adherence to evidence-based standards contributes to higher quality of healthcare.⁴ This assumption, however, seems to lack a clear scientific foundation. Although research suggests that external assessment can have a positive impact on an organisation level, for example, on improved leadership, quality systems and professional development,^{6–9} less is known about the impacts of external assessment on the quality of care. According to a Cochrane review of the literature, there is a paucity of high-quality controlled evaluations on this topic.¹⁰

External assessments are contemporary, real-world events that involve autonomous actors, including healthcare providers, inspecting organs and policy-makers. The complexity of the settings in which external assessments take place may explain why only three randomised controlled studies have been performed to evaluate their effect on quality of care: two small-scale studies^{11 12} and one study¹³ whose methods have been criticised for leading to unreliable conclusions.¹⁰ None of these studies found improvements in patient care resulting from external assessments. Three other studies that used time-series design and a before-and-after design also did not find improvement in performance indicators of care delivery that could be attributed to external assessments.^{14–16} Considering the widespread and growing use of external assessments and the resources spent on conducting and participating in them, there is a need for high-quality studies evaluating their effectiveness on quality of care.¹⁰

The overall aim of this study is to evaluate the effect of statutory inspections at the patient level by assessing detection and treatment of sepsis in emergency departments in Norwegian hospitals. According to WHO's classification, this represents an example of assessment in the supervision category. We use the term "external inspection", in line with the Cochrane review.¹⁰

Sepsis is a major public health challenge and a leading cause of death.¹⁷ The inspections in the present study assessed adherence to standards for sepsis care that have been shown to be associated with improved patient outcomes.^{18 19} Specifically, the aim of our study was to evaluate the effects of the inspections on hospital emergency departments' clinical processes for detecting and treating sepsis and on length of hospital stay and 30-day mortality.

METHODS

To study the effects of the inspections, we used a pragmatic cross-sectional incomplete stepped-wedge design

with cluster-level randomisation at the regional level. The study included data from all hospitals subject to the sepsis inspections.

Setting

The Norwegian healthcare system is publicly funded, and it scores high on Organization for Economic Co-operation and Development (OECD) quality indicators.²⁰ All provision of health services in Norway is regulated by legislation. Healthcare services should be safe, effective and provided in accordance with sound professional practice, and all organisations that provide healthcare services must have a quality management system to ensure that healthcare services are provided in accordance with the legal requirements.²¹ The organisation of specialised emergency care is based on the principle of equal access to services. The government designates a specific geographical area to each hospital, within which the hospital is responsible for providing emergency care to the whole population. Patients within the designated area in need of hospital emergency services are admitted to the hospital after referral or prior contact with general practitioners or other medical professionals.²²

The Norwegian Board of Health Supervision and its regional-level subordinate, the County Governors, are mandated by law to ensure that healthcare is provided in accordance with the legal requirements. An important way of fulfilling this mandate is conducting thematic, nationwide inspections. The themes of these inspections are decided on the basis of information about risk and vulnerability. Norway does not have a mandated system of hospital accreditation, and there are no other regulatory agencies or government bodies supervising the provision of health services.

Intervention

As an intervention, we studied inspections in 24 hospitals in Norway. The inspections addressed early detection and treatment of patients with sepsis admitted to the hospitals' emergency departments. The inspection campaign lasted from April 2016 to March 2018.

The sepsis inspections were planned and directed by the Norwegian Board of Health Supervision, and they were carried out by six inspection teams from the County Governors. Each team performed four inspections within a time frame of about 8 weeks in four geographically proximate hospitals. The inspections were headed by experienced team leaders who were trained in performing inspections. The teams consisted of a minimum of four inspectors with medical and legal expertise, including an independent senior consultant physician in internal medicine or critical care medicine.

The inspections were based on the ISO's procedures for system audits²³ and encompassed three main phases: the announcement of the inspection and collection of relevant data, the site visit, and reporting and follow-up. The information and data reviewed during the inspections comprised administrative documentation, interviews

with management staff and personnel with responsibilities related to care for patients with sepsis, and patient records. During the follow-up period, the inspection teams conducted verification of patient records at 8 and 14 months after the initial inspection.

On the basis of all the information and gathered data, the inspection team assessed whether the emergency department's clinical processes for sepsis detection and treatment were in line with the regulatory standard. A key part of the inspections was identifying and pointing out underlying reasons for substandard performance of the clinical system delivering care to patients admitted with sepsis. The inspection team also assessed to what extent the hospital management had fulfilled their legal obligation to implement a functional management system that monitors, and when necessary improves, the quality of sepsis detection and treatment. In this way, the inspections challenged the quality of performance through addressing the managerial level's responsibility for ensuring good practice through providing an expedient organisational framework for delivering sound professional practice. The inspection teams' findings and conclusions were presented in reports that were made publicly available on the Internet (one of the reports is provided here as an example—see online supplemental material 1). The reports focused on identified non-conformities in the quality of care for patients with sepsis. All inspections found instances of substandard performance, the most common being delay in antibiotic treatment. After each site visit, there was a follow-up phase. During this phase, the hospital management were held responsible for developing and implementing necessary measures to improve substandard performance.

Table 1 provides an overview of the key elements of the intervention.

Study design, participants and data collection

The inspection campaign was mandated to include all 18 counties in Norway, and each County Governor decided which hospitals to inspect in their region. The main inclusion criterion was hospital size because substandard care in larger hospitals would potentially affect more patients. The hospitals selected for inspection comprised all university and regional hospitals and a geographically based selection of local hospitals. They included 24 out of 50 hospitals in Norway with emergency services, and served 75% of the total population.

This study was developed in conjunction with the planning of the inspections. It was not feasible to establish an unexposed control group, and, for practical reasons, the intervention could not be delivered simultaneously to the entire study population. We therefore used a pragmatic cross-sectional incomplete stepped-wedge design with cluster-level randomisation where all inspected hospitals were included in the study.^{24 25} The inspections were carried out sequentially in the 24 hospitals over 12 months from April 2016 to March 2017. The County Governors notified the hospitals of the inspections 2 to 6 months in advance of the site visits. The order of the inspections was randomised at the regional level using a computer-generated list of random numbers; each inspection team received a randomly assigned time slot of 8 weeks during which they were to conduct the inspection of four hospitals in their region. Figure 1 provides an overview of the trial profile. We have previously published the full protocol of the study, including the rationale for the use of the stepped-wedge design.²⁵

We based inclusion into the study on the standard definition of sepsis in use when the study was developed.²⁶ Accordingly, the criteria were clinically suspected infection on presentation to the emergency department and at least two systemic inflammatory response syndrome

Table 1 Key elements of the intervention

Time in months	Activity
1	Inspection team announces inspection and requests the hospital to submit information
2	Inspection team reviews records of patients with sepsis and collect relevant data for the inspection criteria. Data are collected for two time periods, baseline (September 2015) and right before the site visit. Inspection team reviews information from hospital and prepares for the site visit
3	Two-day site visit at the hospital with interviews of key personnel At the end of the site visit, the inspection team presents the preliminary findings, and the hospital can comment on these preliminary findings
4–5	The inspection team writes a preliminary report of their findings. The hospital can comment on the report
6	The inspection team sends the final report to the hospital
Continuously	The hospital plans and implements improvement measures
11	Follow-up audit (8 months after site visit). The inspection team reviews records of patients with sepsis and collect the same kinds of data as they did prior to the site visit Report on findings from audit. Require the hospital to implement necessary changes
17	Follow-up audit (14 months after site visit). The inspection team reviews records of patients with sepsis and collect the same kinds of data as they did prior to the site visit. Report on findings from audit. Require the hospital to implement necessary changes

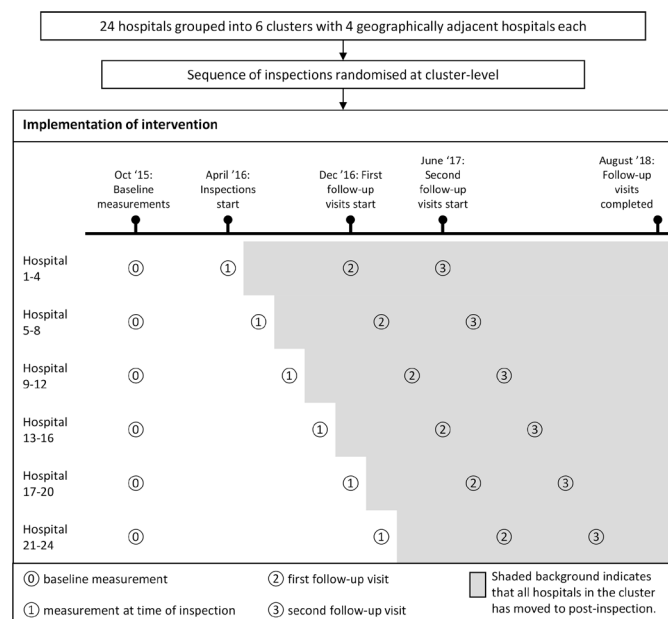


Figure 1 Trial profile.

signs, not including high leucocyte counts. The included patients were aged 18 years or older.

We used a two-step approach to identify eligible patients. First, we searched the Norwegian Patient Registry using a predefined list of the ICD-10 diagnostic codes most commonly used in Norway to classify sepsis and infections (online supplemental material 2).²⁷ The Patient Registry contains diagnostic and therapeutic codes for all hospital admissions. The search produced a list of patients who had been discharged from the participating hospitals with a sepsis and/or infection code, together with an identification number that enabled us to access the corresponding health records. Second, we assessed the individual patient records for eligibility by collecting information about these patients' clinical status on presentation to the emergency department.

We collected data for four time periods for each hospital: two before the inspection and two after the inspection. In the first data collection period, we included patients admitted to all hospitals before 1 October 2015. This was before the Norwegian Board of Health Supervision announced the national inspection campaign. The second collection period varied across the hospitals. The endpoint of this period was the day prior to the site visit at each hospital. The third and fourth time periods were also specific to each hospital, encompassing the 8 and 14 months after the initial site visit, respectively. For each time period, we included the last 83 consecutive patients who fulfilled the inclusion criteria on presentation to the emergency department. For all patient records, we gathered data from the electronic health records about patient age, sex, admission and discharge dates, and the presence of organ failure. Following national evidence-based guidelines,^{26 28} we defined organ failure as fulfilling at least one of the following criteria at arrival to the emergency department:

oxygen saturation $<90\%$ or $\text{PaO}_2/\text{FiO}_2 <40 \text{ kPa}$, altered mental status, urine output $<0.5 \text{ mL/kg/h}$ or increase in serum creatinine $>50 \mu\text{mol/L}$, international normalised ratio >1.5 or activated partial thromboplastin time $>60 \text{ s}$, platelet count <100 or 50% reduction in previous 3 days, serum bilirubin $>70 \text{ mmol/L}$, serum lactate $>4 \text{ mmol/L}$, blood pressure $<90 \text{ mm Hg}$ systolic, mean arterial pressure $<60 \text{ mm Hg}$ or fall in mean arterial pressure $>40 \text{ mm Hg}$. For the first 33 patients, we also collected data on diagnostic measures and treatment given. The data on diagnostic measures and treatment were gathered by the inspection teams and used as audit evidence during the inspection and follow-up visits.

It was not possible to blind health personnel to the intervention, as information about the inspections was publicly known and health personnel participated in interviews and during follow-up. Nor was it possible to blind inspectors and researchers reviewing health records to the intervention or control condition, as information about time and dates was critical to the review.

We obtained data on 30-day all-cause mortality and the Charlson Comorbidity Index²⁹ for the included patients from the Norwegian Patient Registry by connecting data using a unique personal identifier. For patients who had multiple admissions, we used data relating to the first admission.

We performed power calculations using the *stepped-wedge* function³⁰ in Stata/IC, V.14.0 (StataCorp, College Station, TX, USA). For the measures of care delivery, we powered the study to detect an absolute improvement of 70% to 83% and a reduction in mortality of 5% to 11%. We assumed an intra-cluster correlation of 0.05, and type I and type II errors were assumed to be 0.05 and 0.20, respectively. See the study protocol for further details of the power calculations.²⁵

Study outcomes and covariates

Previous research has emphasised the importance of early recognition of sepsis in enabling timely treatment^{31 32} and demonstrated an association between compliance with evidence-based standards and improved outcomes.^{18 19 33} The Norwegian Board of Health Supervision has identified key clinical processes involved in the recognition and treatment of sepsis by examining international guidelines^{26 34} and soliciting advice from experts on sepsis. The indicators we used as study variables for measures of care delivery (see [box 1](#)) were operationalised from the inspection criteria used by the Board of Health Supervision. We defined mortality as all-cause mortality within 30 days of hospital admission, and we defined length of stay as the number of days from admission date to discharge date. In the study protocol, we included the percentage of patients receiving oxygen therapy as a measure of care delivery. However, information on oxygen therapy was not systematically recorded in the electronic patient records, so we had to exclude this variable from our analysis.

Box 1 Study measures

Measures of health care delivery

- ▶ Proportion of patients triaged within 15 min of arrival at the emergency department.*
- ▶ Proportion of patients assessed by a physician in accordance with the urgency specified in the initial triage.
- ▶ Proportion of patients whose vital signs were measured within 1 hour of arrival at the emergency department.
- ▶ Proportion of patients whose blood lactate was measured within 1 hour of arrival at the emergency department.
- ▶ Proportion of patients from whom blood samples[†] were drawn within 1 hour of arrival at the emergency department.
- ▶ Proportion of patients from whom blood cultures were taken before the administration of antibiotics.
- ▶ Proportion of patients with adequate supplementary investigations to detect the focus of infection.
- ▶ Proportion of patients who were adequately observed[‡] while in the emergency department.
- ▶ Proportion of patients who were adequately discharged from the emergency department for further treatment in the hospital (written statement indicating patient status, treatment and further actions).
- ▶ Percentage of patients who received intravenous fluids within 1 hour.
- ▶ Proportion of patients who received antibiotics within 1 hour.

Outcome measures

- ▶ Length of stay in hospital.
- ▶ 30-day all-cause mortality.

*Norwegian hospitals are required to establish a system for prioritising patients admitted to emergency departments. The scales used for this are based on the South African Triage Scale and the Rapid Emergency Triage and Treatment System.

†Leucocyte count, haemoglobin, C reactive protein, creatinine, electrolytes, platelet count and glucose.

‡‘Adequate’ is defined as continuous observation, as well as measurement and documentation of vital signs at least every 15 min in critically ill patients with sepsis and organ failure, measurement and documentation of vital signs every 15 min if a physician has not examined a patient with sepsis but no documented organ failure, and every 30 min after first examination in such patients unless the physician decides otherwise.

Statistical analysis

We compared patient care delivery and outcomes before and after the inspection. Patient characteristics were compared using univariate analyses with Pearson's χ^2 test or the Wilcoxon rank-sum test.

Using logistic regression, we assessed the strength of the associations between having had an inspection and patients receiving adequate care, and we report the associations as ORs. To analyse changes in the patient outcome variables from before the inspection to after the inspection, we used logistic models for 30-day mortality and negative binomial models for length of stay. Because the patient data were sampled from different hospitals, we used mixed-effects models with hospital number included as a random effect to account for clustering.

In our analysis of changes in diagnosis and treatment, we report both unadjusted and adjusted models. In the adjusted models, we controlled for selections of the following variables: age, organ failure, Charlson Comorbidity Index, sex, seasonality and calendar year.

We based the choice of control variables for each model on the Akaike information criteria score. Non-significant adjustment variables were kept in the models if doing so improved the overall model fit (see online supplemental material 3). Age and Charlson Comorbidity Index were entered into the model as linear terms, organ dysfunction and sex were entered as categorical variables, and seasonality and calendar year were entered as categorical variables. For patient outcomes, we report estimates from models with the adjustment variables age, organ dysfunction, Charlson Comorbidity Index and sex, with and without additional adjustment for calendar year.

Among the 3082 health records from which we had collected data on diagnostic and treatment processes, there was missing information on several of the process variables. Data on the diagnostic variables, complete vital measures, blood samples, lactate, observation regimen and supplemental diagnostic procedures were, to a great extent, automatically incorporated in the electronic patient records if they were performed. For these variables, we therefore recoded missing data as the procedure not having been performed within the given time limit. The other variables with missing information were blood culture taken prior to antibiotic treatment (5% missing), timely assessment by a physician (24% missing), time to triage (11% missing), time to fluid administration (29% missing) and time to administration of antibiotics (7% missing). For these variables, we could not assume a specific reason for missing observations. We decided not to impute missing data on these variables because running multiple imputation on dependent variables is not recommended in the absence of an auxiliary variable that correlates strongly with the imputation variable, as was the case here.³⁵

The statistical analyses were performed using Stata/IC, V.16 (StataCorp).

Patient and public involvement

Patient organisations participated in a reference advisory group for the overall research programme, which included this study. They were involved from the planning stage on, but they did not directly participate in developing the research questions or outcome measures used in this article. We used their inputs to inform the overall study design. Patient organisations strongly advocated the importance of disseminating the study findings to relevant parties. The Norwegian Board of Health Supervision has held a national, public conference for hospitals, government agencies and patient representatives where we presented preliminary study findings.

RESULTS

A total of 7407 patients with sepsis were included in the study. Figure 2 shows the flow of patients through the study.

The median age of patients in the pre-inspection and post-inspection groups were 70 and 73 years, respectively.

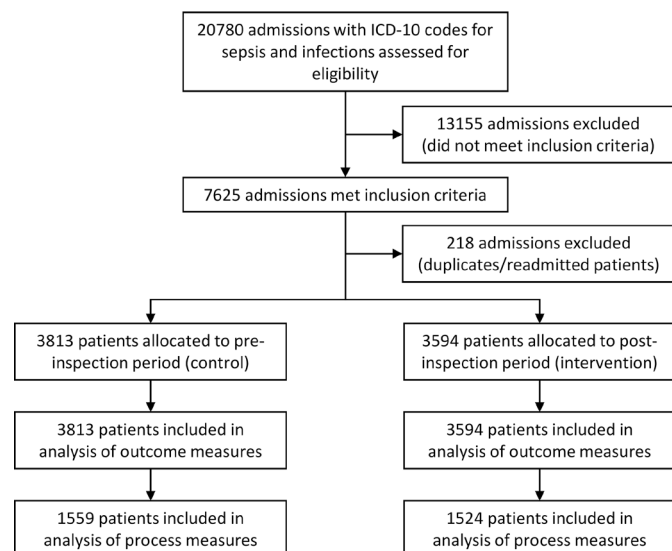


Figure 2 Patient flow.

A larger proportion of patients in the post-inspection group had organ failure (39.3%), compared with the pre-inspection group (36.6%) (see [table 2](#)).

Changes in diagnosis and treatment

Relative to the pre-inspection group, the post-inspection group had higher odds for being examined by a physician within the time frame set in triage (OR 1.28, 95% CI 1.07 to 1.53), higher odds for having complete set of vital measurements taken within 1 hour (OR 1.78, 95% CI 1.10 to 2.87), higher odds for having measured lactate within 1 hour (OR 2.75, 95% CI 1.83 to 4.15), higher odds for having adequate observation regimen (OR 2.20, 95% CI 1.51 to 3.20) and higher odds for antibiotics being administered within 1 hour (OR 2.16, 95% CI 1.83 to 2.55). [Figure 3](#) displays changes in process measures before and after the inspections.

Changes in patient outcomes

On average, the length of stay was significantly shorter in the post-inspection group (6.2 days) than in the pre-inspection group (7.1 days). There was also a significant reduction in all-cause mortality within 30 days of

admission, from 11.1 in the pre-inspection group to 10.5 in the post-inspection group. After controlling for time, neither of these changes were statistically significant (see [table 3](#)).

DISCUSSION

In the adjusted analyses, five of the measures of healthcare delivery showed significant improvement from before to after the inspection. The improvements were observed in aspects of care delivery that are of great importance to patients. We found significant improvement in time to treatment with antibiotics, which is associated with increased survival in patients with severe sepsis and septic shock.^{18 19} Moreover, we also found that significant improvement for timely assessment by a physician, lactate measurement, adequate observation and taking a complete set of vital signs are all key clinical processes in the early detection of sepsis. The first three mentioned measures are significantly associated with earlier treatment with antibiotics.³⁶ The measures that were improved coincided with the main targets of the inspections and the key measures that the hospitals were required to improve following the inspections.

Although there were improvements also in the other measures of healthcare delivery and in the patient outcomes, 30-day mortality and length of stay, these changes were not statistically significant after adjusting for time. This implies that we cannot specifically attribute these improvements to the inspections.

Strengths and limitations

The main strength of our study is that it comprised a robust evaluation of the effects of external inspections on quality of care in real-world settings. To the best of our knowledge, the present research is the largest and most comprehensive study of inspection effects using a cluster-randomised research design.

It is, nevertheless, important to discuss whether the changes in the observed measures are attributable to causes other than the inspections. To do this, we address questions relating to coding and documentation,

Table 2 Patient characteristics

Factor		Before inspection	After inspection	P value
N		3813	3594	
Sex	Male	1939 (50.9%)	1881 (52.3%)	0.2*
	Female	1874 (49.1%)	1713 (47.7%)	
Age, years	Median (IQR)	70 (56–81)	73 (60–82)	<0.001†
	Mean (SD)	66.8 (18.8)	68.9 (18.5)	
Organ failure		1387 (36.6%)	1409 (39.3%)	0.015*
Charlson Comorbidity Index	Median (IQR)	2 (1–4)	2 (1–4)	0.49†

*Pearson's χ^2 test.

†Wilcoxon rank-sum test.

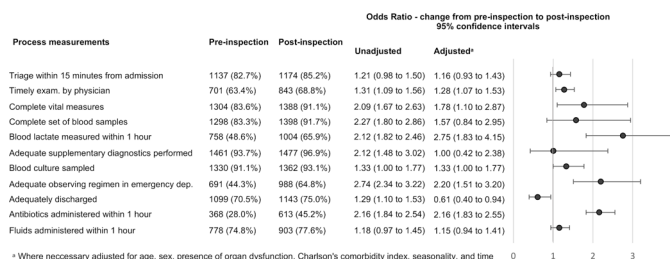


Figure 3 Process measures before and after inspection.

unknown confounders, contamination by other initiatives and time adjustments.

In 2016, the Society of Critical Care Medicine and the European Society of Intensive Care Medicine launched a new sepsis definition.³⁷ This initiative may have led to changes in the coding practice for sepsis. However, inclusion in the present study was based on the assessment of the clinical status of patients with infection at arrival in the emergency department. Patients at Norwegian hospitals are discharged with one main diagnosis and up to seven secondary diagnoses. If a patient has had any kind of infection during the hospital stay, this will be coded as a primary or secondary diagnosis at discharge. Therefore, changes in the coding practice for sepsis could not have biased our analyses. There is a risk that by oversight, no infection or sepsis code was assigned in a health record, resulting in that the record was not included in our screening. However, there is no reason to believe that the relative frequency of such errors would differ between the pre-inspection and post-inspection data. Therefore, such errors are not likely to have influenced our estimates of the inspection effects.

For some of the clinical processes examined in this study, the degree of documentation in the electronic patient records improved after the inspections. To assess whether the observed improvements could be caused by improvements in documentation, we analysed the association between time to treatment and missing data on diagnostic procedures. We found that having missing data on diagnostic procedures was associated with prolonged time to treatment (see online supplemental material 3). Given the association between time to diagnosis and time

to treatment, we expect observations with missing data to be, on average, more delayed, compared with observations for which we have data. Thus, improvements in the process variables are, in all probability, not caused by improvements in documentation.

Due to the preconditions and constraints provided by doing the study in a real-world setting, it was not possible to conduct a randomised controlled trial.²⁵ As with any observational study, there was an inherent risk of confounding from unknown factors. A limitation of our study is that we did not have data on severity of sepsis in the form of commonly used severity scores like SAPS 2 (simplified acute physiology score) or APACHE II (acute physiology and chronic health evaluation), or a detailed organ failure assessment score like SOFA (sequential organ failure assessment). We did control for age, presence of organ failure and comorbidity, which are three important variables associated with severity of sepsis.³⁸

Another potential source of confounding was influence on emergency department practices by other external factors. The stepped-wedge design reduced the risk of such biases. The overlap period where clusters of hospitals switched from control to intervention according to a randomised schedule encompassed 1433 patients. In this period, other factors besides the inspections that could contribute to the observed improvements would affect both the intervention group and the control group simultaneously. In addition, we also controlled for seasonality and year of admittance.

When it comes to confounding by external factors, a particular concern in our study was the possibility of other nationwide initiatives influencing sepsis management at

Table 3 Patient outcomes before and after inspection

	Before	After	Adjusted for background variables*	Adjusted for background and time†
All-cause mortality, 30 days from admission	424 (11.1%)	377 (10.5%)	0.81 (0.69 to 0.95)‡	1.25 (0.86 to 1.80)‡§
Length of stay in days, mean (SD)	7.1 (7.8)	6.2 (6.1)	0.87 (0.84 to 0.90)¶	0.94 (0.86 to 1.03)¶
n	3813	3594	7371	7371

*Adjusted for age, organ dysfunction, sex and Charlson Comorbidity Index.

†Adjusted for secular trends entered as categorical variables per year, plus the other background variables.

‡OR.

§n=7360.

¶Incidence rate ratio.

the emergency departments. There were no other regulatory initiatives that could affect these practices, as the Norwegian Board of Health Supervision and the County Governors are the only bodies assessing compliance with regulatory standards. During the period the present study was conducted, the Norwegian government initiated a voluntary sepsis-management improvement programme. Some of the included hospitals chose to participate in this programme. We specifically analysed the possible impacts of participation in this programme. Participation did not have a significant effect in terms of explaining the improvements in the process measures, and it had a negligible impact on the estimated sizes of the inspection effects. We therefore chose not to include participation in the patient safety programme as a control variable.

‘Anticipatory effects’—stemming from improvement initiatives by the hospitals made in preparation of an upcoming inspection—may also represent a source of bias. Such effects can be considered as constituent parts of the total impact of the inspections.³⁹ Anticipatory initiatives could have influenced the emergency departments’ processes in the timespan between announcement and inspection. As some of the pre-inspection data stems from this period, we could expect our estimates of the inspection effect to be attenuated by the influence of anticipatory effects on the pre-inspection data.

Finally, the time variable may introduce a bias to the regression model which results in an underestimated effect of the intervention. Due to the constraints of doing the research in a real-world setting, the majority of observations were collected from the periods before and after inspection where, respectively, none and all of the hospitals had been inspected. This results in correlation between time and intervention. There are too few observations of inspected hospitals in the early period of the study and of uninspected hospitals in the later period to consistently estimate the true effect of time, independent of the intervention effect. There were significant reductions in mortality and length of stay after the inspections when adjusting for age, organ dysfunction, sex and Charlson Comorbidity Index. When additionally adjusting for time, there were no longer any significant reductions. The lack of significant effects on patient outcome was not expected, as previous research has suggested that the improvements we observed in care delivery would lead to improved patient outcomes,¹⁸ and challenges with modelling of secular trends is one potential explanation for this finding.

Interpretation of findings and comparison with other studies

In contrast to previous studies, which were unable to detect association between external inspections and improvement in quality of care,^{11–13} we found improvements in key measures of care delivery, including time to treatment. The lack of significant associations between inspection and the outcome measures when adjusting for time might be due to the heterogeneity of the patient group included in the study. The effect of earlier

treatment on reduced mortality has been documented in patients with severe sepsis and septic shock.¹⁸ As only a proportion of patients in our study had severe sepsis and septic shock, it seems reasonable to expect more modest reduction in mortality. Furthermore, we need to be careful when reviewing the results from the models where time is included as an adjustment variable. Regardless of whether the improvements in patient outcomes in this study can be attributed to the inspections or not, we found significant improvements in key processes of care delivery including diagnostic processes, observation and time to treatment. These are key processes of care delivery that will enable medical personnel to make sound clinical decisions and initiate treatment processes that have shown to be important to patient outcomes.^{18 36}

The fact that our study showed such improvements after the inspections, whereas previous studies did not, might be explained by contextual factors and how the inspections were conducted. WHO has described a generic framework for how external assessment can contribute to quality improvement: (1) development of standards addressing requirements that will lead to improvement in patient care, (2) reliable identification of performance gaps, (3) involvement of managers and professionals in developing action plans in response to the assessment, and (4) implementation of the plans in a way that lead to improvement.²

Several aspects of the Norwegian sepsis inspections are noteworthy in relation to this framework. Compared with the previously studied inspections,^{15 16} the sepsis inspections in Norway had a narrower target with requirements that were closely related to patient care. Previous work has suggested that, to contribute to improvement, the inspection process should be translated into something meaningful and understandable for clinical practice.^{14 39} The inspections in our study explicitly targeted the early detection and treatment of sepsis, which are crucial for patients and highly relevant and understandable from a clinical perspective. The methods used during the inspections provided the hospitals with reliable and valid data on how their emergency departments performed in terms of the early detection and treatment of patients with sepsis.³⁶ The hospitals were also provided with quantitative and qualitative data that could shed light on possible reasons for substandard performance of the clinical care delivery system.

The inspections addressed the hospital management staff’s responsibility for ensuring evidence-based practice through providing an expedient organisational framework for the clinical care delivery system. During the two follow-up visits, the management was provided with feedback on the progress of the improvement efforts and held accountable for implementing the necessary changes. Previous research has indicated that holding the inspected organisation accountable for implementing changes can contribute to the creation of momentum for implementing the necessary improvement measures.³⁹ Together with orienting the target of the inspections

towards quality of care delivery, this emphasis on following up and holding the hospital management accountable for making changes to the clinical care delivery system may have contributed to the improvements we found.

Adequate discharge, the one process measure which showed a negative trend after the inspections, was not emphasised during the inspections because it was not directly related to early detection of the sepsis diagnosis and treatment. Neither was it a measure that the hospitals were required to report on following the inspection. The fact that this measure was not emphasised during the inspection process might in part explain the negative trend. This finding is also in line with previous research indicating that inspections can have a negative impact on measures that are not within the main purview of the inspections.¹⁵

Our findings also bring into view a larger concern regarding internal and external assessment schemes. There is always a risk that the use of quantitative indicators instigates managerial ‘gaming’, whereby hospitals put efforts into improving their scoring on specific indicators rather than on improving the system delivering care.⁴⁰ The inspections in our study used a set of performance indicators that together provided information about how the clinical system delivering care for patients with sepsis performed as a whole.³⁶ Improving these indicators would thus contribute to an overall systemic improvement in the quality of care for patients with sepsis. If the inspection maintains a clear focus on the overall goal of quality of care by assessing a set of performance indicators that matters for patient care, gaming behaviour by inspected hospitals can be dissuaded.

CONCLUSIONS

Comparing a range of measures before and after inspection, we found improvements in both processes and patient outcomes following the inspections. Though the improvements in patient outcomes cannot be specifically attributed to the inspections, we find substantial improvements in care delivery for patients with sepsis. Our findings indicate that inspections can be used to foster large-scale improvements in quality of care. However, it does not mean that conducting inspections necessarily leads to improvements. Inspections are complex interventions, and one can assume that their efficacy depends on the context in which they are conducted and how they are planned and implemented. Policy-makers and inspecting bodies should therefore prioritise assessing the effects of their inspections and pay attention to the mechanisms by which inspections might contribute to improving care for patients.

Author affiliations

¹Department of Social Science, Western Norway University of Applied Sciences, Sogndal, Norway

²Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway

³Department of Health and Functioning, Western Norway University of Applied Sciences, Bergen, Norway

⁴Clinic of Anesthesia and Intensive Care, St. Olavs Hospital University Hospital in Trondheim, Trondheim, Norway

⁵Gemini Center for Sepsis Research, Department of Circulation and Medical Imaging, Norwegian University of Science and Technology, Trondheim, Norway

⁶Department of Clinical Medicine, University of Bergen, Bergen, Norway

⁷Alliance Manchester Business School, The University of Manchester, Manchester, UK

⁸Institute of Health and Society, University of Oslo, Oslo, Norway

⁹National Centre for Emergency Primary Health Care, NORCE Norwegian Research Centre AS, Bergen, Norway

¹⁰Department of Research, Stavanger University Hospital, Stavanger, Norway

¹¹Department 5, Norwegian Board of Health Supervision, Oslo, Norway

¹²Norwegian Institute of Public Health, Oslo, Norway

¹³Department of Research and Development, Haukeland University Hospital, Bergen, Norway

¹⁴Department of Clinical Science, University of Bergen, Bergen, Norway

Acknowledgements We thank the inspection teams from the County Governors and the clinical experts for participating in data collection. We also thank the staff at the Norwegian Board of Health Supervision who developed the sepsis inspection, and who participated in the data collection and administration of the project. An earlier draft of the manuscript was reviewed by Trish Reynolds and edited by Jennifer Barrett, both from Edanz Group.

Contributors GH: conception, data collection, data curation, statistical analyses, visualisation, project administration, writing – original draft, writing – critical revision, final approval. RMN: conception, statistical analyses, writing – critical revision, final approval. ES: conception, data collection, writing – critical revision, final approval. HKF: conception, writing – critical revision, final approval. KW: conception, writing – critical revision, final approval. JCF: conception, writing – critical revision, final approval. GTB: conception, writing – critical revision, final approval. GSB: conception, data collection, translation of supplemental material, writing – critical revision, final approval. JH: conception, writing – critical revision, final approval. SH: conception, writing – critical revision, final approval. EH: conception, statistical analyses, project administration, supervision, visualisation, writing – original draft, writing – critical revision, final approval.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Disclaimer Data from the Norwegian Patient Registry have been used in this publication. The interpretation and reporting of these data are the sole responsibility of the authors. No endorsement by the Norwegian Patient Registry is intended nor should be inferred.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval Ethics approval was obtained from the Regional Ethics Committee of Norway North (2015/2195/REK nord) and the Norwegian Data Protection Authority (15/01559). The Regional Ethics Committee of Norway Nord ruled that individual patient consent was not required for this study.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. As this study used data from the Norwegian Patient Registry, there are legal restrictions on sharing the data set. The registry is regulated according to the Act relating to Personal Health Data Registries (2014) and specific regulation for the registry under the provision of this act. These laws do not permit us to share these data for secondary use. Third parties can apply to obtain data from the Norwegian Directorate of Health, which is the body granting access to the registry. The Directorate responds to inquiries regarding the Norwegian Patient Registry via email (helseregistre@helsedir.no). Data obtained from the Norwegian Patient Registry can be used to identify electronic health records at the hospitals, pending approval from ethics committee and hospitals.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability

of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Gunnar Husabø <http://orcid.org/0000-0002-8726-1769>

Jon Helgeland <http://orcid.org/0000-0002-9348-7844>

Einar Hovlid <http://orcid.org/0000-0003-0300-4405>

REFERENCES

- Shaw CD, Braithwaite J, Moldovan M, et al. Profiling health-care accreditation organizations: an international survey. *Int J Qual Health Care* 2013;25:222–31.
- Shaw C, Groene O, Berger E. External institutional strategies: accreditation, certification, supervision. In: Busse R, Klazinga N, Panteli D, et al., eds. *Improving healthcare quality in Europe: characteristics, effectiveness and implementation of different strategies*. Copenhagen (Denmark): European Observatory on Health Systems and Policies, 2019.
- Campbell M, Fitzpatrick R, Haines A, et al. Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000;321:694–6.
- Walshe K, Freeman T, Latham L, et al. *Clinical governance – from policy to practice*. Birmingham, UK: University of Birmingham, Health Services Management Centre, 2000.
- Walshe K, Boyd A. *Designing regulation: a review*. Manchester: The University of Manchester – Manchester Business School, 2007.
- Greenfield D, Braithwaite J. Health sector accreditation research: a systematic review. *Int J Qual Health Care* 2008;20:172–83.
- Shaw CD, Groene O, Botje D, et al. The effect of certification and accreditation on quality management in 4 clinical services in 73 European hospitals. *Int J Qual Health Care* 2014;26 Suppl 1:100–7.
- Hinchcliff R, Greenfield D, Moldovan M, et al. Narrative synthesis of health service accreditation literature. *BMJ Qual Saf* 2012;21:979–91.
- Shaw C, Groene O, Mora N, et al. Accreditation and iso certification: do they explain differences in quality management in European hospitals? *Int J Qual Health Care* 2010;22:445–51.
- Flodgren G, Gonçalves-Bradley DC, Pomey M-P. External inspection of compliance with standards for improved healthcare outcomes. *Cochrane Database Syst Rev* 2016;12:CD008992.
- Oude Wesselink SF, Lingsma HF, Ketelaars CAJ, et al. Effects of government supervision on quality of integrated diabetes care: a cluster randomized controlled trial. *Med Care* 2015;53:784–91.
- Oude Wesselink SF, Lingsma HF, Reulings PGJ, et al. Does government supervision improve stop-smoking counseling in midwifery practices? *Nicotine Tob Res* 2015;17:572–9.
- Salmon J, Heavens J, Lombard CBethesda MD, ed. *The impact of accreditation on the quality of hospital care: KwaZulu-Natal Province, Republic of South Africa*. The U.S. Agency for International Development (USAID) by the Quality Assurance Project, University Research Co., LLC, 2003.
- Evaluation team OPM. *Evaluation of the Healthcare Commission's Healthcare Associated Infections Inspection Programme*. London: OPM, 2009.
- Castro-Avila A, Bloor K, Thompson C. The effect of external inspections on safety in acute hospitals in the National health service in England: a controlled interrupted time-series analysis. *J Health Serv Res Policy* 2019;24:1355819619837288:182–90.
- Allen T, Walshe K, Proudlove N, et al. Measurement and improvement of emergency department performance through inspection and rating: an observational study of emergency departments in acute hospitals in England. *Emerg Med J* 2019;36:326–32.
- Rudd KE, Johnson SC, Agesa KM, et al. Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the global burden of disease study. *The Lancet* 2020;395:200–11.
- Levy MM, Rhodes A, Phillips GS, et al. Surviving sepsis campaign: association between performance metrics and outcomes in a 7.5-year study. *Crit Care Med* 2015;43:3–12.
- Seymour CW, Gesten F, Prescott HC, et al. Time to treatment and mortality during mandated emergency care for sepsis. *N Engl J Med* 2017;376:2235–44.
- OECD. *Health at a glance 2017: OECD indicators*. Paris: OECD Publishing, 2017.
- Braut GS. The requirement to practice in accordance with sound professional practice. In: Molven O, Ferkis J, eds. *Healthcare, welfare and law Oslo*. Gyldendal akademisk, 2011.
- Blinkenberg J, Pahlavanyali S, Hetlevik Øystein, et al. General practitioners' and out-of-hours doctors' role as gatekeeper in emergency admissions to somatic hospitals in Norway: registry-based observational study. *BMC Health Serv Res* 2019;19:568.
- ISO. *Conformity Assessment – Requirements and Recommendations for Content of a Third-Party Audit Report on Management Systems*. Geneva: International Organization for Standardization, 2012. ISBN: ISO/IEC TS 17022:2012.
- Hemming K, Lilford R, Girling AJ. Stepped-wedge cluster randomised controlled trials: a generic framework including parallel and multiple-level designs. *Stat Med* 2015;34:181–96.
- Hovlid E, Frich JC, Walshe K, et al. Effects of external inspection on sepsis detection and treatment: a study protocol for a quasiexperimental study with a stepped-wedge design. *BMJ Open* 2017;7:e016213.
- Dellinger RP, Levy MM, Rhodes A, et al. Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock: 2012. *Crit Care Med* 2013;41:580–637.
- Knoop ST, Skrede S, Langeland N, et al. Epidemiology and impact on all-cause mortality of sepsis in Norwegian hospitals: a national retrospective study. *PLoS One* 2017;12:e0187990.
- Norwegian Directorate of Health. *Sepsis. Antibiotika i sykehus*. 2, 2018.
- Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373–83.
- Hemming K, Girling A. A menu-driven facility for power and detectable-difference calculations in stepped-wedge cluster-randomized trials. *Stata J* 2014;14:363–80.
- Gatewood Medley O'Keefe, Wemple M, Greco S, et al. A quality improvement project to improve early sepsis care in the emergency department. *BMJ Qual Saf* 2015;24:787–95.
- Pruinelli L, Westra BL, Yadav P, et al. Delay within the 3-hour surviving sepsis campaign guideline on mortality for patients with severe sepsis and septic shock. *Crit Care Med* 2018;46:500–5.
- Damiani E, Donati A, Serafini G, et al. Effect of performance improvement programs on compliance with sepsis bundles and mortality: a systematic review and meta-analysis of observational studies. *PLoS One* 2015;10:e0125827.
- National Institute for Health and Care Excellence. *Sepsis: recognition, diagnosis and early management*. In: *NICE guideline 51*, 2016. <https://www.nice.org.uk/guidance/ng51>
- Allison P. Missing data. In: *The SAGE Handbook of quantitative methods in psychology*. SAGE Publications Ltd, 2009. <http://methods.sagepub.com/book/sage-hdbk-quantitative-methods-in-psychology>
- Nasa P, Nilsen RM, Flaatten H, et al. Early diagnosis of sepsis in emergency departments, time to treatment, and association with mortality: an observational study. *PLoS One* 2020;15:e0227652.
- Singer M, Deutschman CS, Seymour CW, et al. The third International consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA* 2016;315:801–10.
- Nasa P, Juneja D, Singh O. Severe sepsis and septic shock in the elderly: an overview. *World J Crit Care Med* 2012;1:23–30.
- Smithson R, Richardson E, Roberts J. *Impact of the Care Quality Commission on provider performance: room for improvement?* London: The King's Fund, 2018.
- Bevan G, Hood C. What's measured is what matters: targets and gaming in the english public health care system. *Public Adm* 2006;84:517–38.